

# Finite-time and Asymptotic Convergence of Distributed Averaging and Maximizing Algorithms\*

Guodong Shi and Karl Henrik Johansson<sup>†</sup>

## Abstract

In this paper, we formulate and investigate a generalized consensus algorithm which makes an attempt to unify distributed averaging and maximizing algorithms considered in the literature. Each node iteratively updates its state as a time-varying weighted average of its own state, the minimal state, and the maximal state of its neighbors. We prove that finite-time consensus is almost impossible for averaging under this uniform model. Both time-dependent and state-dependent graphs are considered, and various necessary and/or sufficient conditions are presented on the consensus convergence. For time-dependent graphs, we show that quasi-strong connectivity is critical for averaging, as is strong connectivity for maximizing. For state-dependent graphs defined by a  $\mu$ -nearest-neighbor rule, where each node interacts with its  $\mu$  nearest smaller neighbors and the  $\mu$  nearest larger neighbors, we show that  $\mu + 1$  is a critical threshold on the total number of nodes for the transit from finite-time to asymptotic convergence for averaging, in the absence of node self-confidence. The threshold is  $2\mu$  if each node chooses to connect only to neighbors with unique values. Numerical examples illustrate the tightness of the conditions. The results characterize some fundamental similarities and differences between distributed averaging and maximizing algorithms.

**Keywords:** Averaging algorithms, Max-consensus, Finite-time convergence, State-dependent connections

---

\*This work has been supported in part by the Knut and Alice Wallenberg Foundation, the Swedish Research Council and KTH SRA TNG.

<sup>†</sup>The authors are with ACCESS Linnaeus Centre, School of Electrical Engineering, Royal Institute of Technology, Stockholm 10044, Sweden. Email: {guodongs, kallej}@kth.se

# 1 Introduction

Distributed averaging algorithms, where each node iteratively averages its neighbors' states, have been extensively studied in the literature, due to its wide applicability in engineering [11, 12, 33], computer science [8, 9], and social science [5, 6, 7]. Recently also the max-consensus algorithms have attracted attention. These algorithms compute the maximal value among the nodes, and have been used for leader election, network size estimation, and various applications in wireless networks [33, 31].

The convergence to a consensus is central in the study of averaging and maximizing algorithms but can be hard to analyze, especially when the node interactions are carried out over a switching graph. The most convenient way of modeling the switching node interactions is just to assume the communication graphs are defined by a sequence of time-dependent graphs over the node set. The connectivity of this sequence of graphs plays an important role for the network to reach consensus. Joint connectivity, i.e., connectivity of the union graph over time intervals, has been considered, and various convergence conditions have been established [11, 25, 12, 13, 16, 17, 14, 15, 18, 17]. The switching topology can be dependent on the node states. For instance, in Krause's model, each node is connected only to nodes within a certain distance [23]. Vicsec's model has a similar setting but with higher-order node dynamics [22]. Because the node dynamics is coupled with the graph dynamics for state-dependent graphs, the convergence analysis is quite challenging. Deterministic consensus algorithms with state-dependent graph were studied in [24, 28], and convergence results under probabilistic models were established in [26, 27].

Few studies have discussed the fundamental similarities and differences between distributed averaging and maximizing. Averaging and maximizing consensus algorithms are both distributed information processing over graphs, where nodes communicate and exchange information with its neighbors in the aim of collective convergence. Average consensus algorithms in the literature are based on two standing assumptions: local cohesion and node self-confidence. The node states iteratively update as a weighted average of its neighbors' states, with a positive lower bound for the weight corresponding to its own state [11, 25, 12, 16, 14, 15, 21]. Average consensus algorithms can also be viewed as the equivalent state evolution process where each node updates its state as a weighted average of its own state, and the minimum and maximum states of its neighbors. Maximizing (or minimizing) consensus algorithms are simply based on that each node updates its state to the maximal (minimal) state among its neighbors [32, 35]. Asymptotic

convergence is common in the study of averaging consensus algorithms [14, 16, 21, 12], while it has been shown that maximizing algorithms converge in general in finite time [32, 35]. Finite-time convergence of averaging algorithms was investigated in [33] for a continuous-time model, and recently finite-time consensus in discrete time was discussed in [39] for a special case of gossiping [38].

In this paper, we make the simple observation that averaging and maximizing algorithms can be viewed as instances of a more general distributed processing model. Using this model the transition of the consensus convergence can be studied for the two classes of distributed algorithms in a unified way. Each node iteratively updates its state as a weighted average of its own state together with the minimum and maximum states of its neighbors. By special cases for the weight parameters, averaging and maximizing algorithms can be analyzed. Under this uniform model, we prove for averaging that finite-time consensus is impossible in general, and asymptotical consensus is possible only when the node self-confidence satisfies a divergence condition. Both time-dependent and state-dependent graphs are considered, and various necessary and/or sufficient conditions are presented on the consensus convergence. For time-dependent graphs, we show that quasi-strong connectivity (existence of a spanning tree) is critical for averaging, as is strong connectivity for maximizing. We use a  $\mu$ -nearest-neighbor rule to generate state-dependent graphs, in which each node interacts with its  $\mu$  nearest smaller neighbors ( $\mu$  neighbors with smaller state values), and the nearest  $\mu$  larger neighbors. This model is motivated from recent studies of collective bird behavior [29]. For averaging algorithms without node self-confidence under such state-dependent graphs, we show that  $\mu + 1$  is a critical value for the total number of nodes: finite-time consensus is achieved globally if the number of nodes is no larger than  $\mu + 1$ , and finite-time consensus fails for almost all initial conditions if the number of nodes is larger than  $\mu + 1$ . Moreover, it is shown that this critical number of nodes is instead  $2\mu$  if each node chooses to connect only to neighbors with distinct values in the neighbor rule.

The rest of the paper is organized as follows. In Section 2 we introduce the considered network model, the uniform processing algorithm, and the consensus problem. The impossibilities of finite-time or asymptotic consensus are studied in Section 3. Consensus convergence is studied for time-dependent and state-dependent graphs in Sections 4 and 5, respectively. We show numerical examples in Section 6 and finally some concluding remarks are given in Section 7.

## 2 Problem Definition

In this section, we introduce the network model, the considered algorithm, and define the problem of interest.

### 2.1 Network

We first recall some concepts and notations in graph theory [1]. A directed graph (digraph)  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  consists of a finite set  $\mathcal{V}$  of nodes and an arc set  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . An element  $e = (i, j) \in \mathcal{E}$  is called an *arc* from node  $i \in \mathcal{V}$  to  $j \in \mathcal{V}$ . If the arcs are pairwise distinct in an alternating sequence  $v_0 e_1 v_1 e_2 v_2 \dots e_k v_k$  of nodes  $v_i \in \mathcal{V}$  and arcs  $e_i = (v_{i-1}, v_i) \in \mathcal{E}$  for  $i = 1, 2, \dots, k$ , the sequence is called a (directed) *path* with *length*  $k$ . If there exists a path from node  $i$  to node  $j$ , then node  $j$  is said to be *reachable* from node  $i$ . Each node is thought to be reachable by itself. A node  $v$  from which any other node is reachable is called a *center* (or a *root*) of  $\mathcal{G}$ . A digraph  $\mathcal{G}$  is said to be *strongly connected* if node  $i$  is reachable from  $j$  for any two nodes  $i, j \in \mathcal{V}$ ; *quasi-strongly connected* if  $\mathcal{G}$  has a center [2]. The *distance* from  $i$  to  $j$  in a digraph  $\mathcal{G}$ ,  $d(i, j)$ , is the length of a shortest simple path  $i \rightarrow j$  if  $j$  is reachable from  $i$ , and the *diameter* of  $\mathcal{G}$  is  $\text{diam}(\mathcal{G}) = \max\{d(i, j) | i, j \in \mathcal{V}, j \text{ is reachable from } i\}$ . The union of two digraphs with the same node set  $\mathcal{G}_1 = (\mathcal{V}, \mathcal{E}_1)$  and  $\mathcal{G}_2 = (\mathcal{V}, \mathcal{E}_2)$  is defined as  $\mathcal{G}_1 \cup \mathcal{G}_2 = (\mathcal{V}, \mathcal{E}_1 \cup \mathcal{E}_2)$ . A digraph  $\mathcal{G}$  is said to be *bidirectional* if for every two nodes  $i$  and  $j$ ,  $(i, j) \in \mathcal{E}$  if and only if  $(j, i) \in \mathcal{E}$ . A bidirectional graph  $\mathcal{G}$  is said to be *connected* if there is a path between any two nodes.

Consider a network with node set  $\mathcal{V} = \{1, 2, \dots, n\}$ ,  $n \geq 3$ . Time is slotted. Denote the state of node  $i$  at time  $k \geq 0$  as  $x_i(k) \in \mathbb{R}$ . Then  $x(k) = (x_1(k) \dots x_n(k))^T$  represents the network state. For time-varying graphs, we use the following definition.

**Definition 1 (Time-dependent Graph)** *The interactions among the nodes are determined by a given sequence of digraphs with node set  $\mathcal{V}$ , denoted as  $\mathcal{G}_k = (\mathcal{V}, \mathcal{E}_k)$ ,  $k = 0, 1, \dots$ .*

Throughout this paper, we call node  $j$  a *neighbor* of node  $i$  if there is an arc from  $j$  to  $i$  in the graph. Each node is supposed to always be a neighbor of itself. Let  $\mathcal{N}_i(k)$  represent the neighbor set of node  $i$  at time  $k$ .

## 2.2 Algorithm

The classical average consensus algorithm in the literature is given by

$$x_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) x_j(k), \quad i = 1 \dots, n. \quad (1)$$

Two standing assumptions are fundamental in determining the nature of its dynamics:

**A1** (*Local Cohesion*)  $\sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) = 1$  for all  $i$  and  $k$ ;

**A2** (*Self-confidence*) There exists a constant  $\eta > 0$  such that  $a_{ii}(k) \geq \eta$  for all  $i$  and  $k$ .

These assumptions are widely imposed in the existing works, e.g., [12, 11, 20, 21, 14, 15, 16, 25]. With A1 and A2, we have

$$\sum_{j \in \mathcal{N}_i(k)} a_{ij}(k) x_j(k) = \eta x_i(k) + (a_{ii}(k) - \eta) x_i(k) + \sum_{j \in \mathcal{N}_i(k), j \neq i} a_{ij}(k) x_j(k) \quad (2)$$

and

$$(1 - \eta) \min_{j \in \mathcal{N}_i(k)} x_j(k) \leq (a_{ii}(k) - \eta) x_i(k) + \sum_{j \in \mathcal{N}_i(k), j \neq i} a_{ij}(k) x_j(k) \leq (1 - \eta) \max_{j \in \mathcal{N}_i(k)} x_j(k). \quad (3)$$

Noting the fact that for any  $c \in [a, b]$  there exists a unique  $\lambda \in [0, 1]$  satisfying  $c = \lambda a + (1 - \lambda)b$ , we see from (3) that for every  $i$  and  $k$ , there exists  $\beta_k^{(i)} \in [0, 1]$  such that

$$\begin{aligned} & (a_{ii}(k) - \eta) x_i(k) + \sum_{j \in \mathcal{N}_i(k), j \neq i} a_{ij}(k) x_j(k) \\ &= \beta_k^{(i)} (1 - \eta) \min_{j \in \mathcal{N}_i(k)} x_j(k) + (1 - \beta_k^{(i)}) (1 - \eta) \max_{j \in \mathcal{N}_i(k)} x_j(k) \\ &= \alpha_k^{(i)} \min_{j \in \mathcal{N}_i(k)} x_j(k) + (1 - \eta - \alpha_k^{(i)}) \max_{j \in \mathcal{N}_i(k)} x_j(k), \end{aligned} \quad (4)$$

where  $\alpha_k^{(i)} = \beta_k^{(i)} (1 - \eta) \in [0, 1 - \eta]$ .

Therefore, in light of (2) and (4), based on assumptions A1 and A2, we can always write the average consensus algorithm (1) into the following equivalent form:

$$x_i(k+1) = \eta x_i(k) + \alpha_k^{(i)} \min_{j \in \mathcal{N}_i(k)} x_j(k) + (1 - \eta - \alpha_k^{(i)}) \max_{j \in \mathcal{N}_i(k)} x_j(k), \quad (5)$$

where  $\alpha_k^{(i)} \in [0, 1 - \eta]$  for all  $i$  and  $k$ . Thus, the information processing principle behind distributed averaging is that each node iteratively takes a weighted average of its current state and the minimum and maximum states of its neighbor set.

The standard maximizing algorithm [32, 35] is defined by

$$x_i(k+1) = \max_{j \in \mathcal{N}_i(k)} x_j(k), \quad (6)$$

so distributed maximizing is each node interacting with its neighbors and simply taking the maximal state within its neighbor set.

In this paper, we aim to present a model under which we can discuss fundamental differences of some distributed information processing mechanisms. We consider the following algorithm for the node updates:

$$x_i(k+1) = \eta_k x_i(k) + \alpha_k \min_{j \in \mathcal{N}_i(k)} x_j(k) + (1 - \eta_k - \alpha_k) \max_{j \in \mathcal{N}_i(k)} x_j(k), \quad (7)$$

where  $\alpha_k, \eta_k \geq 0$  and  $\alpha_k + \eta_k \leq 1$ . We denote the set of all algorithms of the form (7) by  $\mathcal{A}$ , when the parameter  $(\alpha_k, \eta_k)$  takes value as  $\eta_k \in [0, 1], \alpha_k \in [0, 1 - \eta_k]$ . This model is a special case of (5) as the parameter  $\alpha_k$  is not depending on the node index  $i$  in (7).

Note that  $\mathcal{A}$  represents a uniform model for distributed averaging and maximizing algorithms. Obeying the cohesion and self-confidence assumptions, the set of (weighted) averaging algorithms,  $\mathcal{A}_{\text{ave}}$ , consists of algorithms in the form of (7) with parameters  $\eta_k \in (0, 1], \alpha_k \in [0, 1 - \eta_k]$ . The set of maximizing algorithms,  $\mathcal{A}_{\text{max}}$ , is given by the parameter set  $\eta_k \equiv 0$  and  $\alpha_k \equiv 0$ .

**Remark 1** *Algorithm (7) is more restrictive than (5) in the sense that the averaging weight  $\alpha_k^{(i)}$  in (5) might vary for different nodes. Hence, (7) cannot in general capture the averaging algorithm (1). Except for this difference, the standing assumptions A1 and A2 of average consensus algorithms are still fulfilled for algorithm (7). Moreover, note that for (7) no lower bound is imposed on the node self-confidence.*

**Remark 2** *In Algorithm (7) each node's update only depends on the states of the minimum and maximum neighbor states at every time step. In other words, not all links are active explicitly in the iterations. Therefore, the existing convergence results on averaging algorithms cannot be applied directly, since these results rely on the connectivity of the communication graph.*

**Remark 3** *If besides A1 we impose a double stochasticity assumption [21, 20, 25] on the arc weights  $a_{ij}(k)$ , i.e.,  $\sum_{i \in \mathcal{N}_j(k)} a_{ji}(k) = 1$  for all  $j$  and  $k$ , each node's state will converge to the average of all initial values during the evolution of Algorithm (1) as long as consensus is reached. In the absence of double stochasticity assumption, if there is a consensus under time-varying communication, we know that the consensus limit of (1) still is a convex combination*

of the initial values with coefficients invariant with respect to initial conditions [14, 15, 21]. However, neither of these conclusions holds for Algorithm (7). It is straightforward to see that the convergence limit is a convex combination of the initial values if consensus is reached. But due to the state-dependent node update in (7), the coefficients in the convex combination of the consensus limit indeed depend on the initial condition (even with fixed communication graph).

### 2.3 Problem

Let  $\{x(k; x^0) = (x_1(k; x^0) \dots x_n(k; x^0))^T\}_0^\infty$  be the sequence generated by (7) for initial time  $k_0$  and initial value  $x^0 = x(k_0) = (x_1(k_0) \dots x_n(k_0))^T \in \mathbb{R}^n$ . We will identify  $x(k; x^0)$  as  $x(k)$  in the following discussions. We introduce the following definition on the convergence of the considered algorithm.

**Definition 2** (i) *Asymptotic consensus is achieved for Algorithm (7) for initial condition  $x(k_0) = x^0 \in \mathbb{R}^n$  if there exists  $z_*(x^0) \in \mathbb{R}$  such that*

$$\lim_{k \rightarrow \infty} x_i(k) = z_*, \quad i = 1, \dots, n.$$

*Global asymptotic consensus is achieved if asymptotic consensus is achieved for all  $k_0 \geq 0$  and  $x^0 \in \mathbb{R}^n$ .*

(ii) *Finite-time consensus is achieved for Algorithm (7) for initial condition  $x(k_0) = x^0 \in \mathbb{R}^n$  if there exist  $z_*(x^0) \in \mathbb{R}$  and an integer  $T_*(x^0) > 0$  such that*

$$x_i(T_*) = z_*, \quad i = 1, \dots, n.$$

*Global finite-time consensus is achieved if finite-time consensus is achieved for all  $k_0 \geq 0$  and  $x^0 \in \mathbb{R}^n$ .*

The algorithm reaching consensus is equivalent with that  $x(k)$  converges to the manifold

$$\mathbf{C} = \left\{ x = (x_1 \dots x_n)^T : x_1 = \dots = x_n \right\}.$$

We call  $\mathbf{C}$  the consensus manifold. Its dimension is one.

In the following, we focus on the impossibilities and possibilities of asymptotic or finite-time consensus. We will show that the convergence properties drastically change when Algorithm (7) transits from averaging to maximizing.

### 3 Convergence Impossibilities

In this section, we discuss the impossibilities of asymptotic or finite-time convergence for the averaging algorithms in  $\mathcal{A}_{\text{ave}}$ . One theorem for each case is proven.

**Theorem 1** *For every averaging algorithm in  $\mathcal{A}_{\text{ave}}$ , finite-time consensus fails for all initial values in  $\mathbb{R}^n$  except for initial values on the consensus manifold.*

*Proof.* We define

$$h(k) = \min_{i \in \mathcal{V}} x_i(k); \quad H(k) = \max_{i \in \mathcal{V}} x_i(k).$$

Introduce  $\Phi(k) = H(k) - h(k)$ . Then clearly asymptotic consensus is achieved if and only if  $\lim_{k \rightarrow \infty} \Phi(k) = 0$ .

Take a node  $i$  satisfying  $x_i(k) = h(k)$ . We have

$$\begin{aligned} x_i(k+1) &= \eta_k x_i(k) + \alpha_k \min_{j \in \mathcal{N}_i(k)} x_j(k) + (1 - \eta_k - \alpha_k) \max_{j \in \mathcal{N}_i(k)} x_j(k) \\ &\leq (\alpha_k + \eta_k) h(k) + (1 - \eta_k - \alpha_k) H(k). \end{aligned} \quad (8)$$

Similarly, taking another node  $m$  satisfying  $x_m(k) = H(k)$ , we obtain

$$\begin{aligned} x_m(k+1) &= \eta_k x_j(k) + \alpha_k \min_{m \in \mathcal{N}_i(k)} x_j(k) + (1 - \eta_k - \alpha_k) \max_{m \in \mathcal{N}_i(k)} x_j(k) \\ &\geq \alpha_k h(k) + (1 - \alpha_k) H(k). \end{aligned} \quad (9)$$

With (8) and (9), we conclude that

$$\Phi(k+1) = \max_{i \in \mathcal{V}} x_i(k) - \min_{i \in \mathcal{V}} x_i(k) \geq x_m(k+1) - x_i(k+1) \geq \eta_k \Phi(k). \quad (10)$$

Therefore, since (10) holds for all  $k$ , we immediately obtain that for every algorithm in the averaging set  $\mathcal{A}_{\text{ave}}$ ,

$$\Phi(K) \geq \Phi(k_0) \prod_{k=k_0}^{K-1} \eta_k > 0 \quad (11)$$

for all  $K \geq k_0$  as long as  $\Phi(k_0) > 0$ . Noticing that the initial values satisfying  $\Phi(k_0) = 0$  are on the consensus manifold, the desired conclusion follows.  $\square$

Since the consensus manifold is a one-dimensional manifold in  $\mathbb{R}^n$ , Theorem 1 indicates that finite-time convergence is almost impossible for average consensus algorithms. This partially explains why finite-time convergence results are rare for averaging algorithms in the literature.

Next, we discuss the impossibility of asymptotic consensus. The following lemma is well-known.



**Lemma 1** *Let  $\{b_k\}_0^\infty$  be a sequence of real numbers with  $b_k \in [0, 1)$  for all  $k$ . Then  $\sum_{k=0}^\infty b_k = \infty$  if and only if  $\prod_{k=0}^\infty (1 - b_k) = 0$ .*

The following theorem on asymptotic convergence holds.

**Theorem 2** *For every averaging algorithm in  $\mathcal{A}_{\text{ave}}$ , asymptotic consensus fails for all initial values in  $\mathbb{R}^n$  except for initial values on the consensus manifold, if  $\sum_{k=0}^\infty (1 - \eta_k) < \infty$ .*

*Proof.* In light of Lemma 1 and (11), we see that for every algorithm in the averaging set  $\mathcal{A}_{\text{ave}}$ ,

$$\lim_{K \rightarrow \infty} \Phi(K) \geq \Phi(k_0) \prod_{k=k_0}^\infty \eta_k > 0 \quad (12)$$

if  $\sum_{k=0}^\infty (1 - \eta_k) < \infty$  for all initial values satisfying  $\Phi(k_0) > 0$ . The desired conclusion thus follows.  $\square$

Theorem 2 indicates that  $\sum_{k=0}^\infty (1 - \eta_k) = \infty$  is a necessary condition for average algorithms to reach asymptotic consensus. Note that  $\eta_k$  measures node self-confidence. Thus, the condition  $\sum_{k=0}^\infty (1 - \eta_k) = \infty$  characterizes the maximal self-confidence that nodes can hold and still reach consensus.

It is worth pointing out that Theorems 1 and 2 hold for any communication graph. In the following discussions, we turn to the possibilities for consensus. Then, however, the communication graph plays an important role. Time-dependent and state-dependent graphs will be studied, respectively, in the following two sections.

## 4 Time-dependent Graphs

In this section, we focus on time-dependent graphs. We first discuss a special case when the network topology is fixed, and the required connectivity for average and max-min algorithms will be treated. Next, time-varying communications will be discussed, and a series of conditions will be presented on the asymptotic or finite-time convergence of the considered algorithm under jointly connected graphs.

### 4.1 Fixed Graph

For fixed communication graphs, we present the following result.

**Theorem 3** *Suppose  $\mathcal{G}_k \equiv \mathcal{G}_*$  is a fixed digraph for all  $k$ .*

(i) For every algorithm in  $\mathcal{A}_{\text{ave}}$ , global asymptotic consensus can be achieved only if  $\mathcal{G}_*$  is quasi-strongly connected.

(ii) For every algorithm in  $\mathcal{A}_{\text{max}}$ , global finite-time consensus is achieved if and only if  $\mathcal{G}_*$  is strongly connected.

*Proof.* (i) If  $\mathcal{G}_*$  is not quasi-strongly connected, there exist two distinct nodes  $i$  and  $j$  such that  $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$ , where  $\mathcal{V}_1 = \{\text{nodes from which } i \text{ is reachable in } \mathcal{G}_*\}$  and  $\mathcal{V}_2 = \{\text{nodes from which } j \text{ is reachable in } \mathcal{G}_*\}$ . Consequently, nodes in  $\mathcal{V}_1$  never receive information from nodes in  $\mathcal{V}_2$ . Take  $x_i(k_0) = 0$  for  $i \in \mathcal{V}_1$  and  $x_i(k_0) = 1$  for  $i \in \mathcal{V}_2$ . Obviously, consensus cannot be achieved under this initial condition. The conclusion holds.

(ii) (Sufficiency.) Let  $v_0$  be a node with the maximal value initially. Then after one step all the nodes for which  $v_0$  is a neighbor will reach the maximal value. Proceeding the analysis we see that the whole network will converge to the initial maximum in finite time.

(Necessity.) Assume that  $\mathcal{G}_*$  is not strongly connected. There will be two different nodes  $i_*$  and  $j_*$  such that  $j_*$  is not reachable from  $i_*$ . Introduce  $\mathcal{V}_* = \{j : j \text{ is reachable from } i_*\}$ . Then  $\mathcal{V}_* \neq \mathcal{V}$  because  $j_* \notin \mathcal{V}_*$ . Moreover, the definition of  $\mathcal{V}_*$  guarantees that all the nodes in  $\mathcal{V} \setminus \mathcal{V}_*$  will never be influenced by the nodes in  $\mathcal{V}_*$ . Therefore, letting the initial maximum be unique and reached by some node in  $\mathcal{V}_*$ , consensus will not be reached.

The proof is complete.  $\square$

As will be shown in the following discussions, quasi-strong connectivity is not only necessary, but also sufficient to guarantee global asymptotic consensus for the algorithms in the averaging set  $\mathcal{A}_{\text{ave}}$  under some mild conditions on the parameters  $(\alpha_k, \eta_k)$ . Therefore, Theorem 3 clearly states that quasi-strong connectivity is critical for averaging, as is strong connectivity for maximizing.

## 4.2 Time-varying Graph

We now turn to time-varying graphs. Denote the joint graph of  $\mathcal{G}_k$  over time interval  $[k_1, k_2]$  as  $\mathcal{G}([k_1, k_2]) = (\mathcal{V}, \cup_{k \in [k_1, k_2]} \mathcal{E}(k))$ , where  $0 \leq k_1 \leq k_2 \leq +\infty$ . We introduce the following definitions on the joint connectivity of time-varying graphs.

**Definition 3** (i)  $\mathcal{G}_k$  is uniformly jointly quasi-strongly connected (strongly connected) if there exists an integer  $B \geq 1$  such that  $\mathcal{G}([k, k+B-1])$  is quasi-strongly connected (strongly connected) for all  $k \geq 0$ .

- (ii)  $\mathcal{G}_k$  is infinitely jointly strongly connected if  $\mathcal{G}([k, \infty))$  is strongly connected for all  $k \geq 0$ .
- (iii) Suppose  $\mathcal{G}_k$  is bidirectional for all  $k$ . Then  $\mathcal{G}_k$  is infinitely jointly connected if  $\mathcal{G}([k, \infty))$  is connected for all  $k \geq 0$ .

**Remark 4** The uniformly joint connectivity, which requires the union graph to be connected over each bounded interval, has been extensively studied in the literature, e.g., [11, 12, 14, 15, 16]. The infinitely joint connectivity is a more general case since it does not impose an upper bound for the length of the interval where connectivity is guaranteed for the union graph. Convergence results for consensus algorithms based on infinitely joint connectivity are given in [17, 18, 19].

The following conclusion holds for uniformly jointly quasi-strongly connected graphs.

**Theorem 4** Suppose  $\mathcal{G}_k$  is uniformly jointly quasi-strongly connected. Algorithms in the averaging set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if either

$$\sum_{s=0}^{\infty} \left[ \prod_{k=s(n-1)^2 B}^{(s+1)(n-1)^2 B-1} \alpha_k \right] = \infty \quad (13)$$

or

$$\sum_{s=0}^{\infty} \left[ \prod_{k=s(n-1)^2 B}^{(s+1)(n-1)^2 B-1} (1 - \alpha_k - \eta_k) \right] = \infty. \quad (14)$$

Theorem 4 hence states that divergence of certain products of the algorithm parameters guarantees global asymptotic consensus.

It is straightforward to see that for a non-negative sequence  $\{b_k\}$  with  $b_k \geq b_{k+1}$  for all  $k$ ,  $\sum_{s=0}^{\infty} \prod_{k=s(n-1)^2 B}^{(s+1)(n-1)^2 B-1} b_k = \infty$  if and only if  $\sum_{k=0}^{\infty} b_k^{(n-1)^2 B} = \infty$ . Thus, the following corollary follows from Theorem 4.

**Corollary 1** Suppose  $\mathcal{G}_k$  is uniformly jointly quasi-strongly connected.

- (i) Assume that  $\alpha_k \geq \alpha_{k+1}$  for all  $k$ . Algorithms in the averaging set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if  $\sum_{k=0}^{\infty} \alpha_k^{(n-1)^2 B} = \infty$ .
- (ii) Assume that  $\alpha_k + \eta_k \leq \alpha_{k+1} + \eta_{k+1}$  for all  $k$ . Algorithms in the averaging set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if  $\sum_{k=0}^{\infty} (1 - \alpha_k - \eta_k)^{(n-1)^2 B} = \infty$ .

For uniformly jointly strongly connected graphs, it turns out that consensus can be achieved under weaker conditions on  $(\alpha_k, \eta_k)$ .

**Theorem 5** Suppose  $\mathcal{G}_k$  is uniformly jointly strongly connected. Algorithms in the averaging set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if either

$$\sum_{s=0}^{\infty} \left[ \prod_{k=s(n-1)B}^{(s+1)(n-1)B-1} \alpha_k \right] = \infty \quad (15)$$

or

$$\sum_{s=0}^{\infty} \left[ \prod_{k=s(n-1)B}^{(s+1)(n-1)B-1} (1 - \alpha_k - \eta_k) \right] = \infty. \quad (16)$$

Similarly, Theorem 5 leads to the following corollary.

**Corollary 2** Suppose  $\mathcal{G}_k$  is uniformly jointly strongly connected.

(i) Assume that  $\alpha_k \geq \alpha_{k+1}$  for all  $k$ . Averaging algorithms in the set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if  $\sum_{k=0}^{\infty} \alpha_k^{(n-1)B} = \infty$ .

(ii) Assume that  $\alpha_k + \eta_k \leq \alpha_{k+1} + \eta_{k+1}$  for all  $k$ . Averaging algorithms in the set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if  $\sum_{k=0}^{\infty} (1 - \alpha_k - \eta_k)^{(n-1)B} = \infty$ .

For bidirectional graphs, the conditions are much simpler to state. We present the following result.

**Theorem 6** Suppose  $\mathcal{G}_k$  is bidirectional for all  $k$  and  $\mathcal{G}_k$  is infinitely jointly connected. Averaging algorithms in the set  $\mathcal{A}_{\text{ave}}$  achieve global asymptotic consensus if there exists a constant  $\alpha_* \in (0, 1)$  such that either  $\alpha_k \geq \alpha_*$  or  $1 - \alpha_k - \eta_k \geq \alpha_*$  for all  $k$ .

The convergence of algorithms in the maximizing set  $\mathcal{A}_{\text{max}}$  is stated as follows.

**Theorem 7** Maximizing algorithms in the set  $\mathcal{A}_{\text{max}}$  achieve global finite-time consensus if  $\mathcal{G}_k$  is infinitely jointly strongly connected.

Theorems 4–7 together provide a comprehensive understanding of the convergence conditions for the considered model (7) under time-varying graphs. Infinitely jointly strong connectivity is sufficient for global finite-time consensus for algorithms in  $\mathcal{A}_{\text{max}}$  according to Theorem 7, while infinitely joint connectivity cannot ensure global asymptotic consensus for algorithms in  $\mathcal{A}_{\text{ave}}$  in general. Thus, in this sense algorithms in  $\mathcal{A}_{\text{ave}}$  and  $\mathcal{A}_{\text{max}}$  are fundamentally different under infinitely jointly connected graphs.

The rest of this section contains the proofs of Theorems 4–7.

### 4.2.1 Proof of Theorem 4

We continue to use the following notations:

$$h(k) = \min_{i \in \mathcal{V}} x_i(k), \quad H(k) = \max_{i \in \mathcal{V}} x_i(k),$$

and  $\Phi(k) = H(k) - h(k)$ . Following any solution of (7), it is obvious to see that  $h(k)$  is non-decreasing and  $H(k)$  is non-increasing.

Note that if (13) guarantees asymptotic consensus for algorithm (7), replacing the node states  $x_i(k)$  with  $-x_i(k)$  leads to that (14) guarantees asymptotic consensus of algorithm (7) for  $-x_i(k), i = 1, \dots, n$ . Since consensus for  $x_i(k), i = 1, \dots, n$  is equivalent with consensus for  $-x_i(k), i = 1, \dots, n$ , (13) and (14) are equivalent in terms of consensus convergence. Thus, we just need to show that (13) is a sufficient condition for asymptotic consensus.

Take  $k_* \geq 0$  as any moment in the iterative algorithm. Take  $(n-1)^2$  intervals  $[k_*, k_* + B - 1]$ ,  $[k_* + B, k_* + 2B - 1], \dots, [k_* + (n^2 - 2n)B, k_* + (n-1)^2B - 1]$ . Since  $\mathcal{G}_k$  is uniformly jointly quasi-strongly connected, the union graph on each of these intervals has at least one center node. Consequently, there must be a node  $v_0$  and  $n-1$  integers  $0 \leq b_1 < b_2 < \dots < b_{n-1} \leq n^2 - 2n$  such that  $v_0$  is a center of  $\mathcal{G}([k_* + b_i B, k_* + (b_i + 1)B - 1]), i = 1, \dots, n-1$ . Assume that

$$x_{v_0}(k_*) \leq \frac{1}{2}h(k_*) + \frac{1}{2}H(k_*). \quad (17)$$

We first bound  $x_{v_0}(k)$  for  $k \in [k_*, k_* + (n-1)^2B]$ . It is not hard to see that

$$\begin{aligned} x_{v_0}(k_* + 1) &= \eta_{k_*} x_{v_0}(k_*) + \alpha_{k_*} \min_{j \in \mathcal{N}_{v_0}(k_*)} x_j(k_*) + (1 - \alpha_{k_*} - \eta_{k_*}) \max_{j \in \mathcal{N}_{v_0}(k_*)} x_j(k_*) \\ &\leq (\alpha_{k_*} + \eta_{k_*}) \left( \frac{1}{2}h(k_*) + \frac{1}{2}H(k_*) \right) + (1 - \alpha_{k_*} - \eta_{k_*}) H(k_*) \\ &\leq \alpha_{k_*} \left( \frac{1}{2}h(k_*) + \frac{1}{2}H(k_*) \right) + (1 - \alpha_{k_*}) H(k_*) \\ &= \frac{\alpha_{k_*}}{2} h(k_*) + \left( 1 - \frac{\alpha_{k_*}}{2} \right) H(k_*). \end{aligned} \quad (18)$$

Proceeding, we obtain

$$x_{v_0}(k_* + m) \leq \frac{\prod_{k=k_*}^{k_*+m-1} \alpha_k}{2} h(k_*) + \left( 1 - \frac{\prod_{k=k_*}^{k_*+m-1} \alpha_k}{2} \right) H(k_*), \quad m = 0, 1, \dots \quad (19)$$

Since  $v_0$  is a center of  $\mathcal{G}([k_* + b_1 B, k_* + (b_1 + 1)B - 1])$ , there exists another node  $v_1$  such that  $v_0$  is a neighbor of  $v_1$  for some  $k_1 \in [k_* + b_1 B, k_* + (b_1 + 1)B - 1]$ . As a result, based on

(19), we have

$$\begin{aligned}
x_{v_1}(k_1 + 1) &= \eta_{k_1} x_{v_1}(k_1) + \alpha_{k_1} \min_{j \in \mathcal{N}_{v_1}(k_1)} x_j(k_1) + (1 - \alpha_{k_1} - \eta_{k_1}) \max_{j \in \mathcal{N}_{v_1}(k_1)} x_j(k_1) \\
&\leq \alpha_{k_1} x_{v_0}(k_1) + (1 - \alpha_{k_1}) H(k_1) \\
&\leq \alpha_{k_1} \left( \frac{\prod_{k=k_*}^{k_1-1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_1-1} \alpha_k}{2}\right) H(k_*) \right) + (1 - \alpha_{k_1}) H(k_*) \\
&= \frac{\prod_{k=k_*}^{k_1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_1} \alpha_k}{2}\right) H(k_*). \tag{20}
\end{aligned}$$

Proceeding, we have

$$x_{v_0}(k_* + m) \leq \frac{\prod_{k=k_*}^{k_*+m-1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_*+m-1} \alpha_k}{2}\right) H(k_*), \quad m = (b_1 + 1)B, \dots \tag{21}$$

Continuing the analysis on time intervals  $[k_* + b_i B, k_* + (b_i + 1)B - 1]$  for  $i = 2, \dots, n - 1$  and nodes  $v_2, v_3, \dots, v_{n-1}$ , similar upper bounds for each node can be obtained:

$$x_{v_i}(k_* + m) \leq \frac{\prod_{k=k_*}^{k_*+m-1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_*+m-1} \alpha_k}{2}\right) H(k_*), \quad m = (b_i + 1)B, \dots \tag{22}$$

This immediately leads to

$$x_{v_i}(k_* + (n - 1)^2 B) \leq \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2}\right) H(k_*), \quad i = 0, 1, \dots, n \tag{23}$$

which implies

$$H(k_* + (n - 1)^2 B) \leq \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2}\right) H(k_*). \tag{24}$$

Thus, we have

$$\begin{aligned}
\Phi(k_* + (n - 1)^2 B) &= H(k_* + (n - 1)^2 B) - h(k_* + (n - 1)^2 B) \\
&\leq \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2} h(k_*) + \left(1 - \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2}\right) H(k_*) - h(k_*) \\
&= \left(1 - \frac{\prod_{k=k_*}^{k_*+(n-1)^2 B-1} \alpha_k}{2}\right) \Phi(k_*). \tag{25}
\end{aligned}$$

From a symmetric analysis by bounding  $h(k_* + (n - 1)^2 B)$  from below, we know that (25) also holds for the other condition with  $x_{v_0}(k_*) \geq \frac{1}{2}h(k_*) + \frac{1}{2}H(k_*)$ . Therefore, since  $k_*$  is selected arbitrarily, we can assume the initial time is  $k_0 = 0$ , without loss of generality, and then conclude that

$$\Phi(K(n - 1)^2 B) \leq \Phi(0) \prod_{s=0}^{K-1} \left(1 - \frac{1}{2} \prod_{k=s(n-1)^2 B}^{(s+1)(n-1)^2 B-1} \alpha_k\right). \tag{26}$$

The desired conclusion follows immediately from Lemma 1.

#### 4.2.2 Proof of Theorem 5

Notice that in a strongly connected graph, every node is a center node. Therefore, when  $\mathcal{G}_k$  is uniformly jointly strongly connected, taking  $k_* \geq 0$  as any moment in the iteration and  $n - 1$  intervals  $[k_*, k_* + B - 1]$ ,  $[k_* + B, k_* + 2B - 1]$ ,  $\dots$ ,  $[k_* + (n - 2)B, k_* + (n - 1)B - 1]$ , any node  $i \in \mathcal{V}$  is a center node for the union graph over each of these intervals. As a result, the desired conclusion follows repeating the analysis used in the proof of Theorem 4.

#### 4.2.3 Proof of Theorem 6

Similar to the proof of Theorem 4, we only need to show that the existence of a constant  $\alpha_* \in (0, 1)$  such that  $\alpha_k \geq \alpha_*$  is sufficient for asymptotic consensus.

Take  $k_1^* \geq 0$  as an arbitrary moment in the iterative algorithm. Take a node  $u_0 \in \mathcal{V}$  satisfying  $x_{u_0}(k_1^*) = h(k_1^*)$ . We define

$$k_1 = \inf \{k \geq k_1^* : \text{there exists another node connecting } u_0 \text{ at time } k\} \quad (27)$$

and then

$$\mathcal{V}_1 = \{k \geq k_1^* : \text{nodes which are connected to } u_0 \text{ at time } k_1\}. \quad (28)$$

Thus, we have

$$\begin{aligned} x_{u_0}(k_1 + 1) &= \eta_{k_1} x_{u_0}(k_1) + \alpha_{k_1} \min_{j \in \mathcal{N}_{u_0}(k_1)} x_j(k_1) + (1 - \alpha_{k_1} - \eta_{k_1}) \max_{j \in \mathcal{N}_{u_0}(k_1)} x_j(k_1) \\ &\leq (\alpha_{k_1} + \eta_{k_1}) x_{u_0}(k_1) + (1 - \alpha_{k_1} - \eta_{k_1}) H(k_1) \\ &\leq (\alpha_{k_1} + \eta_{k_1}) h(k_1^*) + (1 - \alpha_{k_1} - \eta_{k_1}) H(k_1^*) \\ &\leq \alpha_{k_1} h(k_1^*) + (1 - \alpha_{k_1}) H(k_1^*) \\ &\leq \alpha_* h(k_1^*) + (1 - \alpha_*) H(k_1^*) \end{aligned} \quad (29)$$

and

$$\begin{aligned} x_i(k_1 + 1) &= \eta_{k_1} x_i(k_1) + \alpha_{k_1} \min_{j \in \mathcal{N}_i(k_1)} x_j(k_1) + (1 - \alpha_{k_1}) \max_{j \in \mathcal{N}_i(k_1)} x_j(k_1) \\ &\leq \alpha_{k_1} x_{u_0}(k_1) + (1 - \alpha_{k_1}) H(k_1) \\ &\leq \alpha_{k_1} h(k_1^*) + (1 - \alpha_{k_1}) H(k_1^*) \\ &\leq \alpha_* h(k_1^*) + (1 - \alpha_*) H(k_1^*) \end{aligned} \quad (30)$$

for all  $i \in \mathcal{V}_1$ .

Note that if nodes in  $\{u_0\} \cup \mathcal{V}_1$  are not connected with other nodes in  $\mathcal{V} \setminus (\{u_0\} \cup \mathcal{V}_1)$  during  $[k_1 + 1, k_1 + s]$ ,  $s \geq 1$ , we have that for all  $i \in \{u_0\} \cup \mathcal{V}_1$ ,

$$x_i(k_1 + m) \leq \alpha_* h(k_1^*) + (1 - \alpha_*) H(k_1^*), \quad m = 1, \dots, s + 1. \quad (31)$$

Continuing the estimate,  $k_2, \dots, k_d$  and  $\mathcal{V}_2, \dots, \mathcal{V}_d$  can be defined correspondingly until  $\mathcal{V} = \{u_0\} \cup (\cup_{i=1}^d \mathcal{V}_i)$ , so eventually we have

$$x_i(k_d + 1) \leq \alpha_*^d h(k_1^*) + (1 - \alpha_*^d) H(k_1^*), \quad i = 1, \dots, n, \quad (32)$$

which implies

$$H(k_d + 1) \leq \alpha_*^d h(k_1^*) + (1 - \alpha_*^d) H(k_1^*). \quad (33)$$

We denote  $k_2^* = k_d + 1$ . Because it holds that  $d \leq n - 1$ , we see from (33) that

$$\Phi(k_2^*) \leq (1 - \alpha_*^{n-1}) \Phi(k_1^*). \quad (34)$$

Since  $\mathcal{G}_k$  is infinitely jointly connected, this process can be carried on for an infinite sequence  $k_1^* < k_2^* < \dots$ . Thus, asymptotic consensus is achieved for all initial conditions. This completes the proof.

#### 4.2.4 Proof of Theorem 7

Let  $v_0$  be a node with the maximal value initially. Because  $\mathcal{G}_k$  is infinitely jointly strongly connected, we can define

$$k_1 = \inf \{k \geq k_1^* : \text{there exists another node for which } v_0 \text{ is a neighbor at time } k\} \quad (35)$$

and then

$$\mathcal{V}_1 = \{k \geq k_1^* : \text{nodes for which } v_0 \text{ is a neighbor at time } k_1\}. \quad (36)$$

Then at time  $k_1 + 1$  all the nodes in  $\mathcal{V}_1$  will reach the maximal value. Proceeding the analysis we know that the whole network will converge to the initial maximum in finite time.  $\square$



## 5 State-dependent Graphs

In this section, we investigate the convergence of Algorithm (7) for state-dependent graphs. We are interested in a particular set of averaging algorithms,  $\mathcal{A}_{\text{ave}}^*$ , where  $(\alpha_k, \eta_k)$  takes value  $\eta_k \equiv 0$ ,  $\alpha_k \in (0, 1)$ . Algorithms in  $\mathcal{A}_{\text{ave}}^*$  correspond to the case when the self-confidence assumption A2 does not hold, and are of the form

$$x_i(k+1) = \alpha_k \min_{j \in \mathcal{N}_i(k)} x_j(k) + (1 - \alpha_k) \max_{j \in \mathcal{N}_i(k)} x_j(k). \quad (37)$$

Algorithms in  $\mathcal{A}_{\text{ave}}^*$  still have local cohesion. Hence, they fulfill Assumption A1 but not A2. In fact, averaging algorithms without self-confidence have been investigated in classical works on the convergence of product of stochastic matrices, e.g., [3, 4, 5].

### 5.1 State-dependent Communication

In both Krause's [23] and Vicsek's [22] models, nodes interact with neighbors whose distance is within a certain communication range. Convergence analysis for consensus algorithms under such models can be found in [24, 28, 26, 27]. Recently, it was discovered through empirical data that in a bird flock each bird seems to interact with a fixed number of nearest neighbors, rather than with all neighbors within a fixed metric distance [29]. Nearest-neighbor model has been studied under a probabilistic setting on the graph connectivity for wireless communication networks [30]. From a social network point of view, the evolution of opinions may result from similar models since members tend to exchange information with a fixed number of other members who hold a similar opinion as themselves [5, 28]. In this section, we consider a network model in which nodes interact only with other nodes having a close state value. Consider the following nearest-neighbor rule.

**Definition 4 (Nearest-neighbor Graph)** *For a positive integer  $\mu$  and any node  $i \in \mathcal{V}$ , there is a link entering  $i$  from each node in the set  $\mathcal{N}_i^-(k) \cup \mathcal{N}_i^+(k)$ , where*

$$\mathcal{N}_i^-(k) = \{\text{nearest } \mu \text{ neighbors from } \{j \in \mathcal{V} : x_j(k) < x_i(k)\}\}$$

*denotes the nearest smaller neighbor set, and*

$$\mathcal{N}_i^+(k) = \{\text{nearest } \mu \text{ neighbors from } \{j \in \mathcal{V} : x_j(k) > x_i(k)\}\}$$

*denotes the nearest larger neighbor set. The graph defined by this nearest neighbor rule is denoted as  $\mathcal{G}_{x(k)}^\mu$ ,  $k = 0, 1, \dots$ .*

Naturally, if there are less than  $\mu$  nodes with states smaller than  $x_i(k)$ ,  $\mathcal{N}_i^-(k)$  has less than  $\mu$  elements. Similar condition holds for  $\mathcal{N}_i^+(k)$ . Hence, the number of neighbor nodes is not necessarily fixed in the nearest-neighbor graph.

**Remark 5** *Note that, at each time  $k$ , the nearest-neighbor graph is uniquely determined by the node states. The node interactions are indeed determined by the distance between the node states. In this sense, the nearest-neighbor graph shares similar structure with Krause's model [23, 24], where each node communicates with the nodes within certain radius. This nearest-neighbor graph also fulfills the interaction structure in the bird flock model discussed in [29] since each node communicates with an almost fixed number of neighbors, nearest from above and below.*

Note that in the definition of the nearest-neighbor graph, nodes may have neighbors with the same state values. We consider the following nearest-value graph, where each node considers only neighbors with different state values.

**Definition 5** (*Nearest-value Graph*) *For a positive integer  $\mu$  and any node  $i \in \mathcal{V}$ , there is a link entering  $i$  from each node in the set  $\mathcal{N}_i^-(k) \cup \mathcal{N}_i^+(k)$ , where*

$$\mathcal{N}_i^-(k) = \{\text{nearest } \mu \text{ neighbors with different values from } \{j \in \mathcal{V} : x_j(k) < x_i(k)\}\}$$

*denotes the nearest smaller neighbor set, and*

$$\mathcal{N}_i^+(k) = \{\text{nearest } \mu \text{ neighbors with different values from } \{j \in \mathcal{V} : x_j(k) > x_i(k)\}\}$$

*denotes the nearest larger neighbor set. The graph defined by this nearest neighbor rule is denoted as  $\mathcal{G}_{x(k)}^{\mu v}$ ,  $k = 0, 1, \dots$ .*

An illustration of nearest-neighbor and nearest-value graphs at a specific time instance  $k$  is shown in Figure 1 for  $n = 4$  nodes and  $\mu = 2$ .

## 5.2 Basic Lemmas

We first establish two useful lemmas for the analysis of nearest-neighbor and nearest-value graphs. The following lemma indicates that the order of node states is preserved.

**Lemma 2** *For any two nodes  $u, v \in \mathcal{V}$  and every algorithm in  $\mathcal{A}$ , under either the nearest-neighbor graph  $\mathcal{G}_{x(k)}^\mu$  or the nearest-value graph  $\mathcal{G}_{x(k)}^{\mu v}$ , we have*

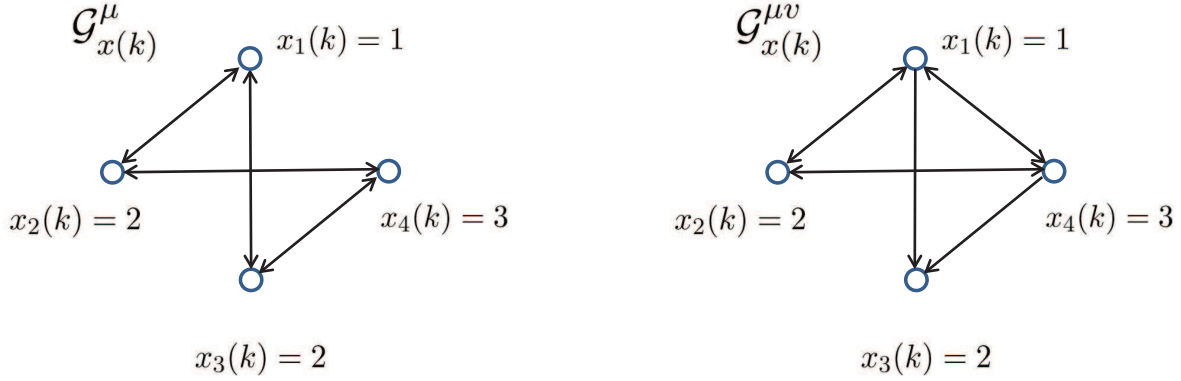


Figure 1: Examples of nearest-neighbor graph  $\mathcal{G}_{x(k)}^\mu$  and nearest-value graph  $\mathcal{G}_{x(k)}^{\mu v}$  for  $\mu = 2$ . Note that for a given set of states, these graphs are in general not unique.

- (i)  $x_u(k+1) = x_v(k+1)$  if  $x_u(k) = x_v(k)$ ;
- (ii)  $x_u(k+1) \leq x_v(k+1)$  if  $x_u(k) < x_v(k)$ .

*Proof.* When  $x_u(k) = x_v(k)$ , we have  $\{j : x_j(k) < x_u(k)\} = \{j : x_j(k) < x_v(k)\}$  and  $\{j : x_j(k) > x_u(k)\} = \{j : x_j(k) > x_v(k)\}$ . Thus, for either  $\mathcal{G}_{x(k)}^\mu$  or  $\mathcal{G}_{x(k)}^{\mu v}$ , both

$$\min_{j \in \mathcal{N}_u(k)} x_j(k) = \min_{j \in \mathcal{N}_v(k)} x_j(k) \quad \text{and} \quad \max_{j \in \mathcal{N}_u(k)} x_j(k) = \max_{j \in \mathcal{N}_v(k)} x_j(k)$$

hold. Then (i) follows straightforwardly.

If  $x_u(k) < x_v(k)$ , it is easy to see that

$$\min_{j \in \mathcal{N}_u(k)} x_j(k) \leq \min_{j \in \mathcal{N}_v(k)} x_j(k); \quad \max_{j \in \mathcal{N}_u(k)} x_j(k) \leq \max_{j \in \mathcal{N}_v(k)} x_j(k)$$

according to the definition of neighbor sets, which implies (ii) immediately.  $\square$

Define

$$\Upsilon_k = \left| \{x_1(k), \dots, x_n(k)\} \right|$$

as the number of distinct node states at time  $k$ , where  $|S|$  for a set  $S$  represents its cardinality. Then Lemma 2 implies that  $\Upsilon_{k+1} \leq \Upsilon_k$  for all  $k \geq 0$ . This point plays an important role in the convergence analysis.

Moreover, for both the nearest-neighbor graph  $\mathcal{G}_{x(k)}^\mu$  and the nearest-value graph  $\mathcal{G}_{x(k)}^{\mu v}$ , in order to distinguish the node states under different values of neighbors, we denote  $x_i^\mu(k)$  as the state of node  $i$  when the number of larger or smaller neighbors is  $\mu$ . Correspondingly, we denote

$$h^\mu(k) = \min_{i \in \mathcal{V}} x_i^\mu(k), \quad H^\mu(k) = \max_{i \in \mathcal{V}} x_i^\mu(k).$$

and  $\Phi^\mu(k) = H^\mu(k) - h^\mu(k)$ . We give another lemma indicating that the convergence speed increases as the number of neighbors increases, which is quite intuitive because apparently graph connectivity increases as the number of neighbors increases.

**Lemma 3** *Consider either the nearest-neighbor graph  $\mathcal{G}_{x(k)}^\mu$  or the nearest-value graph  $\mathcal{G}_{x(k)}^{\mu\nu}$ . Given two integers  $1 \leq \mu_1 \leq \mu_2$ . For every algorithm in  $\mathcal{A}$  and every initial value, we have  $\Phi^{\mu_1}(k) \geq \Phi^{\mu_2}(k)$  for all  $k$ .*

*Proof.* Fix the initial condition at time  $k_0$ . Let  $m \in \mathcal{V}$  be a node satisfying  $x_m^{\mu_1}(k_0) = h^{\mu_1}(k_0)$  and  $x_m^{\mu_2}(k_0) = h^{\mu_2}(k_0)$ . The order preservation property given by Lemma 2 guarantees that  $x_m^{\mu_1}(k) = h^{\mu_1}(k)$  and  $x_m^{\mu_2}(k) = h^{\mu_2}(k)$  for all  $k \geq k_0$ . It is straightforward to see that  $x_m^{\mu_1}(k_0+1) \leq x_m^{\mu_2}(k_0+1)$  if  $\mu_1 \leq \mu_2$ , and continuing we know that  $x_m^{\mu_1}(k_0+s) \leq x_m^{\mu_2}(k_0+s)$  for all  $s \geq 2$ . Thus, we have  $h^{\mu_1}(k) \leq h^{\mu_2}(k)$  for all  $k \geq k_0$ . A symmetric analysis leads to  $H^{\mu_1}(k) \geq H^{\mu_2}(k)$  for all  $k$  and the desired conclusion thus follows.  $\square$

### 5.3 Convergence for Nearest-neighbor Graph

For algorithms in the set  $\mathcal{A}_{\text{ave}}^*$ , we present the following result under nearest-neighbor graph.

**Theorem 8** *Consider the nearest-neighbor graph  $\mathcal{G}_{x(k)}^\mu$ .*

- (i) *When  $n \leq \mu + 1$ , each algorithm in  $\mathcal{A}_{\text{ave}}^*$  achieves global finite-time consensus;*
- (ii) *When  $n > \mu + 1$ , each algorithm in  $\mathcal{A}_{\text{ave}}^*$  fails to achieve finite-time consensus for almost all initial values;*
- (iii) *When  $n > \mu + 1$ , each algorithm in  $\mathcal{A}_{\text{ave}}^*$  achieves global asymptotic consensus if  $\{\alpha_k\}$  is monotone.*

*Proof.* (i) When  $n \leq \mu + 1$ , the communication graph is the complete graph. Thus, consensus will be achieved in one step following (7) for every algorithm in  $\mathcal{A}_{\text{ave}}^*$ .

(ii) Let  $n > \mu + 1$ . We define two index set

$$\mathcal{I}_k^- = \{i : x_i(k) = h(k) = \min_{i \in \mathcal{V}} x_i(k)\}; \quad \mathcal{I}_k^+ = \{i : x_i(k) = H(k) = \max_{i \in \mathcal{V}} x_i(k)\}.$$

*Claim.* Suppose both  $\mathcal{I}_k^-$  and  $\mathcal{I}_k^+$  contain one node only. Then so do  $\mathcal{I}_{k+1}^-$  and  $\mathcal{I}_{k+1}^+$ .

Let  $u$  and  $v$  be the unique element in  $\mathcal{I}_k^-$  and  $\mathcal{I}_k^+$ , respectively. Take  $m \in \mathcal{V} \setminus \{u\}$ . Noting the fact that  $x_m(k) > x_u(k)$  and  $\mu \leq n - 2$ , we have

$$\min_{j \in \mathcal{N}_u(k)} x_j(k) \leq \min_{j \in \mathcal{N}_m(k)} x_j(k), \quad \max_{j \in \mathcal{N}_u(k)} x_j(k) < \max_{j \in \mathcal{N}_m(k)} x_j(k).$$

This leads to  $x_m(k+1) > x_u(k+1)$ . Therefore,  $u$  is still the unique element in  $\mathcal{I}_{k+1}^-$ . Similarly we can prove that  $v$  is still the unique element in  $\mathcal{I}_{k+1}^+$ . The claim holds.

Now observe that

$$\Delta \doteq \bigcup_{u \neq v} \left\{ x = (x_1 \dots x_n)^T : x_u < \min_{m \in \mathcal{V} \setminus \{u\}} x_m \text{ and } x_v > \max_{m \in \mathcal{V} \setminus \{v\}} x_m \right\}$$

has measure zero with respect to the standard Lebesgue measure on  $\mathbb{R}^n$ . For any initial value not in  $\Delta$ , we have both  $\mathcal{I}_k^-$  and  $\mathcal{I}_k^+$  contain one unique element, and thus finite-time consensus is impossible. The desired conclusion follows.

(iii) Recall that

$$\Upsilon_k = \left| \{x_1(k), \dots, x_n(k)\} \right|.$$

Since  $\Upsilon_{k+1} \leq \Upsilon_k$  holds for all  $k$  according to Lemma 2, there exists two integers  $0 \leq m \leq n$  and  $T \geq 0$  such that

$$\Upsilon_k = m, \tag{38}$$

for all  $k \geq T$ . Thus, we can sort the possible node states for all  $k \geq T$  as

$$y_1(k) < y_2(k) < \dots < y_m(k).$$

Apparently  $m \neq 1, 2$  since otherwise the graph is complete for time  $\ell$  with  $\Upsilon_\ell = 1, 2$  and consensus is reached after one step. We assume  $m \geq 3$  in the following discussions.

Algorithm (37) can be equivalently transformed to the dynamics on  $\{y_1(k), \dots, y_m(k)\}$ . Moreover, based on Lemma 3, we only need to prove asymptotic consensus for the case  $\mu = 1$ .

Let  $\mu = 1$  and  $k \geq T$ . For algorithms in  $\mathcal{A}_{\text{ave}}^*$ , the dynamics of  $\{y_1(k), \dots, y_m(k)\}$  can be written:

$$\begin{cases} y_1(k+1) = \alpha_k y_1(k) + (1 - \alpha_k) y_2(k); \\ y_2(k+1) = \alpha_k y_1(k) + (1 - \alpha_k) y_3(k); \\ \vdots \\ y_{m-1}(k+1) = \alpha_k y_{m-2}(k) + (1 - \alpha_k) y_m(k); \\ y_m(k+1) = \alpha_k y_{m-1}(k) + (1 - \alpha_k) y_m(k). \end{cases} \tag{39}$$

Now let  $\{\alpha_k\}$  be monotone, say, non-decreasing. Then we have  $\alpha_k \geq \alpha_T > 0$ . Therefore, for all  $k \geq T$ , we have

$$y_1(k+1) = \alpha_k y_1(k) + (1 - \alpha_k) y_2(k) \leq \alpha_T y_1(k) + (1 - \alpha_T) y_m(k), \tag{40}$$

and continuing we know that

$$y_1(k+s) \leq \alpha_T^s y_1(k) + (1 - \alpha_T^s) y_m(k), s \geq 1. \quad (41)$$

Similarly for  $y_2(k)$ , we have

$$\begin{aligned} y_2(k+2) &= \alpha_{k+1} y_1(k+1) + (1 - \alpha_{k+1}) y_3(k+1) \\ &\leq \alpha_T (\alpha_T y_1(k) + (1 - \alpha_T) y_m(k)) + (1 - \alpha_T) y_m(k) \\ &\leq \alpha_T^2 y_1(k) + (1 - \alpha_T^2) y_m(k) \end{aligned} \quad (42)$$

and

$$y_2(k+s) \leq \alpha_T^s y_1(k) + (1 - \alpha_T^s) y_m(k), s \geq 2. \quad (43)$$

Proceeding the analysis, eventually we arrive at

$$y_i(k+n-1) \leq \alpha_T^{n-1} y_1(k) + (1 - \alpha_T^{n-1}) y_m(k), i = 1, \dots, n, \quad (44)$$

which yields

$$\begin{aligned} \Phi(k+n-1) &= y_m(k+n-1) - y_1(k+n-1) \\ &\leq \alpha_T^{n-1} y_1(k) + (1 - \alpha_T^{n-1}) y_m(k) - y_1(k) \\ &= (1 - \alpha_T^{n-1}) \Phi(k). \end{aligned} \quad (45)$$

Thus, global asymptotic consensus is achieved. The other case with  $\{\alpha_k\}$  being non-increasing can be proved using a symmetric argument. The desired conclusion follows.

This completes the proof of the theorem.  $\square$

**Remark 6** In Theorem 8, the asymptotic consensus statement relies on the condition that  $\{\alpha_k\}$  is monotone. From the proof of Theorem 8 we see that this condition can be replaced by that there exists a constant  $\varepsilon \in (0, 1)$  such that either  $\alpha_k \geq \varepsilon$  or  $\alpha_k \leq 1 - \varepsilon$  for all  $k$ . In fact, we conjecture that the asymptotic consensus statement of Theorem 8 holds true for all  $\{\alpha_k\}$ , i.e., we believe that asymptotic consensus is achieved for all algorithms in  $\mathcal{A}_{\text{ave}}^*$  under nearest-neighbor graphs.

**Remark 7** Theorem 8 indicates that  $\mu + 1$  is a critical number of nodes for nearest-neighbor graphs: for algorithms in  $\mathcal{A}_{\text{ave}}^*$ , finite-time consensus holds globally if  $n \leq \mu + 1$ , and fails almost globally if  $n > \mu + 1$ . Note that  $n \leq \mu + 1$  implies that the communication graph is the complete

graph, which is rare in general. Recalling that Theorem 1 showed that finite-time consensus fails almost globally for algorithms in  $\mathcal{A}_{\text{ave}}$ , we conclude that finite-time consensus is generally rare for averaging algorithms in  $\mathcal{A}$ , no matter with  $(\mathcal{A}_{\text{ave}})$  or without  $(\mathcal{A}_{\text{ave}}^*)$  the self-confidence assumption.

For algorithms in  $\mathcal{A}_{\text{max}}$ , we present the following result.

**Theorem 9** *Consider the nearest-neighbor graph  $\mathcal{G}_{x(k)}^\mu$ . Algorithms in  $\mathcal{A}_{\text{max}}$  achieve global finite-time consensus in no more than  $\lceil \frac{n}{\mu} \rceil$  steps, where  $\lceil z \rceil$  represents the smallest integer no smaller than  $z$ .*

*Proof.* Without loss of generality, we assume that  $x_1(0), \dots, x_n(0)$  are mutually different. We sort the initial values of the nodes as

$$x_{i_1}(0) < x_{i_2}(0) < \dots < x_{i_n}(0).$$

Here  $i_m$  denotes node with the  $m$ 'th largest value initially.

Observing that  $i_n$  is a right-hand side neighbor of nodes  $i_{n-\mu}, i_{n-\mu+1}, \dots, i_{n-1}$ , we have

$$x_{i_\tau}(1) = x_{i_n}(0), \tau = n - \mu, \dots, n.$$

This leads to  $\Upsilon_1 = \Upsilon_0 - \mu$  where  $\Upsilon_k = |\{x_1(k), \dots, x_n(k)\}|$ . Proceeding the same analysis we know that consensus is achieved in no more than  $\lceil \frac{n}{\mu} \rceil$  steps. The desired conclusion follows.  $\square$

## 5.4 Convergence for Nearest-value Graph

In this subsection, we study the convergence for nearest-value graphs. Since nearest-value graph  $\mathcal{G}_{x(k)}^{\mu\nu}$  indeed increases the connectivity of  $\mathcal{G}_{x(k)}^\mu$ , the asymptotic consensus statement of Theorem 8 also holds for  $\mathcal{G}_{x(k)}^{\mu\nu}$ . The main result for finite-time consensus of nearest-value graphs is presented as follows. It turns out that the critical number of nodes for nearest-value graphs is  $2\mu$ .

**Theorem 10** *Consider the nearest-value graph  $\mathcal{G}_{x(k)}^{\mu\nu}$ .*

- (i) *When  $n \leq 2\mu$ , algorithms in  $\mathcal{A}_{\text{ave}}^*$  achieve global finite-time consensus in no more than  $\lceil \log_2(2\mu + 1) \rceil$  steps;*
- (ii) *When  $n > 2\mu$ , algorithms in  $\mathcal{A}_{\text{ave}}^*$  fail to achieve finite-time consensus for almost all initial conditions.*

*Proof.* (i) Suppose  $n \leq 2\mu$ . Based on Lemma 3, without loss of generality, we assume  $n = 2\mu$  and the initial values of the nodes are mutually different. Now we have  $\Upsilon_0 = |\{x_1(0), \dots, x_n(0)\}| = 2\mu$ . We first show the following claim.

*Claim.* If  $\Upsilon_k = 2\mu - A$  with  $A \geq 0$  an integer, then  $\Upsilon_{k+1} \leq \Upsilon_k - A - 1$ .

We order the node states at time  $k$  and denote them as

$$Y_1 < Y_2 < \dots < Y_{\Upsilon_k}.$$

When  $\Upsilon_k = 2\mu - A$ , it is not hard to find that for all  $m = \mu - A, \dots, \mu + 1$ , each node with value  $Y_{\Upsilon_m}$  will connect to some node with value  $Y_1$ , and some other node with value  $Y_{\Upsilon_k}$ . Therefore, the nodes with value  $Y_{\Upsilon_m}, m = \mu - A, \dots, \mu + 1$  will reach the same state after the  $k$ 'th update. The claim holds.

Therefore, by induction we have  $\Upsilon_k = \max\{0, \Upsilon_0 - \sum_{m=0}^{k-1} 2^m\} = \max\{0, 2\mu - (2^k - 1)\}$ . The conclusion (i) follows straightforwardly.

(ii) Suppose  $n > 2\mu$ . Let  $x_1(0), \dots, x_n(0)$  be mutually different. Then it is not hard to see that for any two nodes  $u$  and  $v$  with  $x_u(0) < x_v(0)$ , at least one of

$$\min_{j \in \mathcal{N}_u(0)} x_j(0) < \min_{j \in \mathcal{N}_v(0)} x_j(0)$$

or

$$\max_{j \in \mathcal{N}_u(0)} x_j(0) < \max_{j \in \mathcal{N}_v(0)} x_j(0)$$

holds. This immediately leads to  $x_u(1) < x_v(1)$ . Because  $u$  and  $v$  are arbitrarily chosen, we can conclude that  $\Upsilon_1 = \Upsilon_0$ . By an induction argument we see that  $\Upsilon_k = \Upsilon_0 = n$  for all  $k \geq 0$ , or equivalently, consensus cannot be achieved in finite time. Now observe that

$$\bigcup_{i \geq j} \{x = (x_1 \dots x_n)^T : x_i = x_j\}$$

has measure zero with respect to the standard Lebesgue measure on  $\mathbb{R}^n$ . The desired conclusion thus follows.  $\square$

## 6 Numerical Examples

In this section, we present two numerical examples.

**Example 1.** Consider a network with six nodes  $\mathcal{V} = \{1, \dots, 6\}$ . The communication graph is fixed and directed, as indicated in Figure 2. Initial values for each node are  $x_i(0) = i - 1, i = 1, \dots, 6$ .



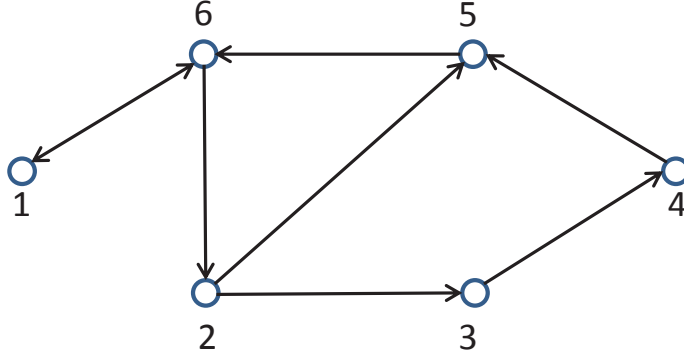


Figure 2: Directed communication graph.

Let  $\eta_k \equiv 0$  and  $\alpha_k \equiv \alpha$  in Algorithm (7). Consider eleven values in the interval  $[0, 1/2]$  of the parameter  $\alpha$ . Recall that  $\Phi(k) = \max_{i \in \mathcal{V}} x_i(k) - \min_{i \in \mathcal{V}} x_i(k)$  is the consensus measure. The trajectory corresponding to each value of  $\alpha$  is shown in Figure 3. We see that finite-time convergence is achieved only when  $\alpha$  is zero, i.e., only for the maximizing case of Algorithm (7).

**Example 2.** Consider a network with  $n = 128$  nodes. Take initial value for the  $i$ 'th node as  $x_i(0) = i$ . Let  $\eta_k \equiv 0$  and  $\alpha_k \equiv 0.5$  in Algorithm (7). The communication graph is given by the nearest-neighbor or nearest-value rule. The evolution of the convergence properties depending on  $\mu$  is shown in Figure 4 and 5. In Figure 4, we illustrate the conclusion in Theorem 8 about that finite-time convergence does not hold under nearest-neighbor graphs  $\mathcal{G}_{x(k)}^\mu$  for  $\mu < 127$ . In Figure 5, we illustrate the conclusion in Theorem 10 about that finite-time convergence does not hold under nearest-value graphs  $\mathcal{G}_{x(k)}^{\mu v}$  for  $\mu < 64$ .

In Figure 6, we show the dependence of the convergence rates on the number of neighbors for nearest-value graphs. We see that there is a sharp increase of the convergence rate when  $\mu$  is larger than about five. This is consistent with the convergence speed for consensus algorithm on Cayley graphs, where the convergence rate increases very fast as the graph degree increases [37].

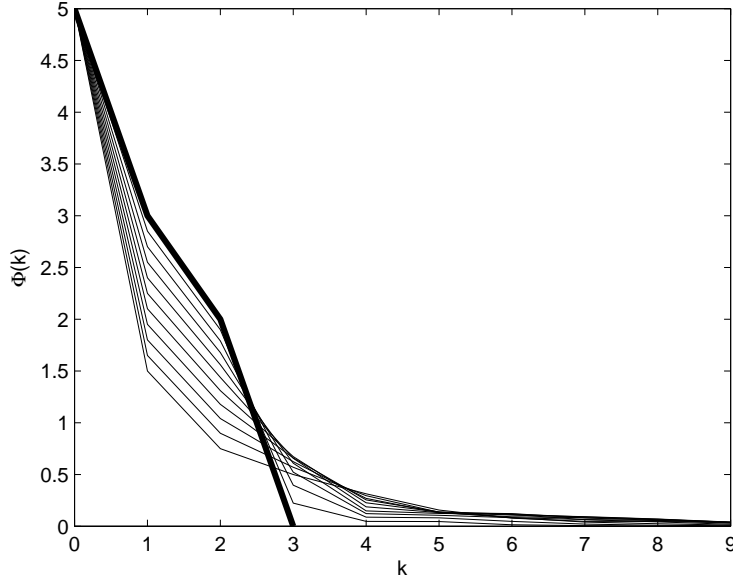


Figure 3: Convergence of Algorithm (7) for  $\alpha = 0, 0.05, 0.1, \dots, 0.5$ . The bold line indicates the trajectory with  $\alpha = 0$ , which has finite-time convergence. All other cases have asymptotic convergence.

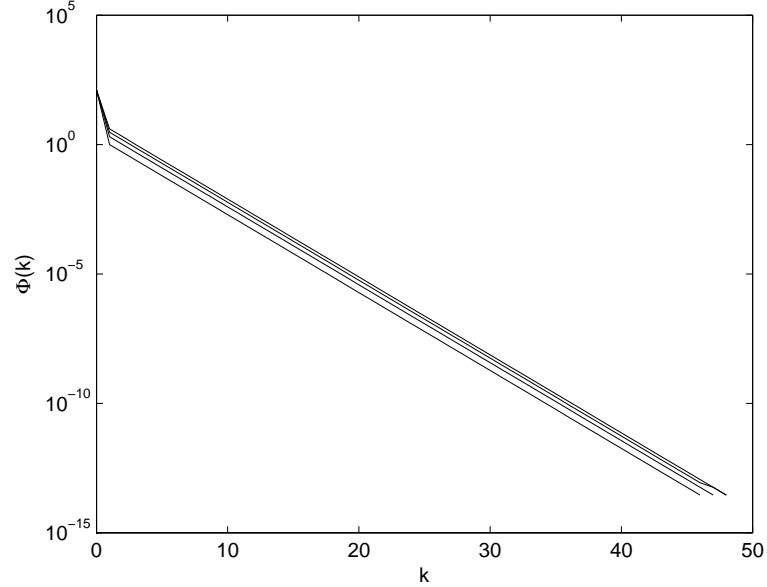


Figure 4: The consensus measure  $\Phi$  for nearest-neighbor graph with  $\mu = 123, \dots, 126$  (upper curve to lower curve). Finite-time convergence does not hold for  $\mu < 127$  according to Theorem 8.

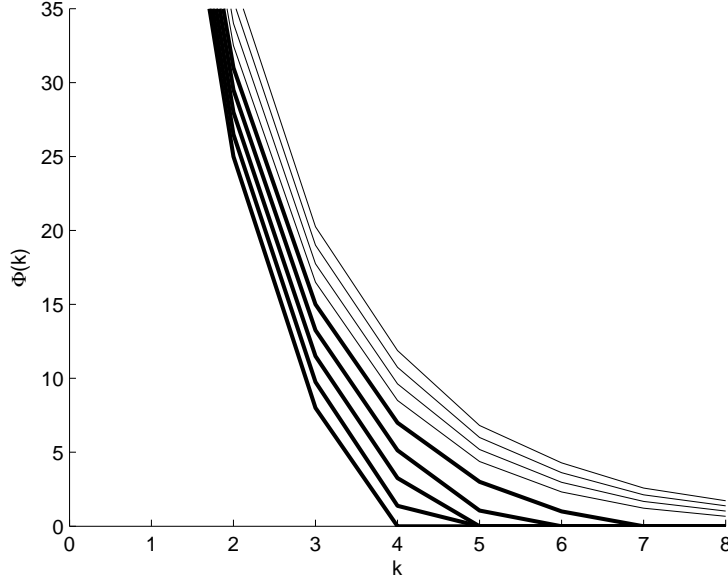


Figure 5: The consensus measure  $\Phi$  for nearest-value graph with  $\mu = 60, \dots, 68$  (upper curve to lower curve). The bold curves ( $\mu = 64, \dots, 68$ ) indicate the cases with finite-time convergence. Finite-time convergence does not hold for  $\mu < 64$ . The simulations illustrate Theorem 10.

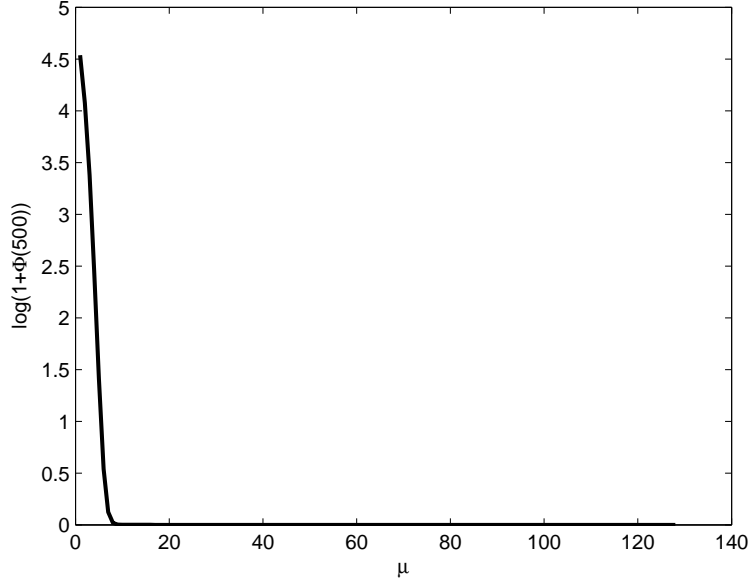


Figure 6: The convergence speed is heavily influenced by  $\mu$  for nearest-value graphs as illustrated in the plot by the value of  $\Phi(k)$  after  $k = 500$  steps. There is a sharp increase of the convergence rate when  $\mu$  increases above 5.

## 7 Conclusions

This paper focused on a uniform model for distributed averaging and maximizing. Each node iteratively updated its state as a weighted average of its own state, the minimal state, and maximal state among its neighbors. We proved that finite-time consensus is almost impossible for averaging under the uniform model. The communication graph was time-dependent or state-dependent. Necessary and sufficient conditions were established on the graph to ensure a global consensus. For time-dependent graphs, we showed that quasi-strong connectivity is critical for averaging algorithms, as is strong connectivity for maximizing algorithms. For state-dependent graphs defined by a  $\mu$ -nearest-neighbor rule, where each node interacts with its  $\mu$  nearest smaller neighbors and the  $\mu$  nearest larger neighbors, we showed that  $\mu + 1$  is a critical number of nodes when consensus transits from finite time to asymptotic convergence in the absence of node self-confidence: finite-time consensus disappears suddenly when the number of nodes is larger than  $\mu + 1$ . This critical number of nodes turned out to be  $2\mu$  if each node chooses to connect to nodes with different values. The results revealed the fundamental connection and difference between distributed averaging and maximizing, but more challenges still lie in the principles underlying the two types of algorithms, such as their convergence rates.

## Acknowledgment

The authors would like to thank Dr. Kin Cheong Sou for his generous help on the numerical examples.

## References

- [1] C. Godsil and G. Royle. *Algebraic Graph Theory*. New York: Springer-Verlag, 2001.
- [2] C. Berge and A. Ghouila-Houri. *Programming, Games, and Transportation Networks*. John Wiley and Sons, New York, 1965.
- [3] J. Wolfowitz, “Products of indecomposable, aperiodic, stochastic matrices,” *Proc. Amer. Math. Soc.*, vol. 15, pp. 733-736, 1963.
- [4] J. Hajnal, “Weak ergodicity in non-homogeneous Markov chains,” *Proc. Cambridge Philos. Soc.*, no. 54, pp. 233-246, 1958.

- [5] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118-121, 1974.
- [6] P. M. DeMarzo, D. Vayanos, J. Zwiebel, "Persuasion bias, social influence, and unidimensional opinions," *Quarterly Journal of Economics*, vol. 118, no. 3, pp. 909-968, 2003.
- [7] B. Golub and M. O. Jackson, "Naïve learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, vol. 2, no. 1, pp. 112-149, 2007.
- [8] S. Muthukrishnan, B. Ghosh, and M. Schultz, "First and second order diffusive methods for rapid, coarse, distributed load balancing," *Theory of Computing Systems*, vol. 31, pp. 331-354, 1998.
- [9] R. Diekmann, A. Frommer, and B. Monien, "Efficient schemes for nearest neighbor load balancing," *Parallel Computing*, vol. 25, pp. 789-812, 1999.
- [10] N. Malpani, J. L. Welch and N. Vaidya, "Leader election algorithms for mobile ad hoc networks," *Proc. 4th international workshop on Discrete algorithms and methods for mobile computing and communications*, pp. 96-104, 2000.
- [11] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Trans. Autom. Control*, vol. 31, pp. 803-812, 1986.
- [12] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 988-1001, 2003.
- [13] R. Olfati-Saber and R. Murray, "Consensus problems in the networks of agents with switching topology and time delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520-1533, 2004.
- [14] M. Cao, A. S. Morse and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: a graphical approach," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 575-600, 2008.
- [15] M. Cao, A. S. Morse and B. D. O. Anderson, "Reaching a consensus in a dynamically changing environment: convergence rates, measurement delays, and asynchronous events," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 601-623, 2008.

- [16] W. Ren and R. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 655-661, 2005.
- [17] L. Moreau, "Stability of multi-agent systems with time-dependent communication links," *IEEE Trans. Autom. Control*, vol. 50, pp. 169-182, 2005.
- [18] G. Shi and Y. Hong, "Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies," *Automatica*, vol. 45, no. 5, pp. 1165-1175, 2009.
- [19] G. Shi, Y. Hong and K. H. Johansson, "Connectivity and set tracking of multi-agent systems guided by multiple moving leaders," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 663-676, 2012.
- [20] A. Nedić, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Trans. Autom. Control*, vol. 54, no. 11, pp. 2506-2517, 2009.
- [21] A. Olshevsky and J. N. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Review*, vol. 53, no. 4, pp. 747-772, 2011.
- [22] T. Vicsek, A. Czirok, E. B. Jacob, I. Cohen, and O. Schochet. Novel type of phase transitions in a system of self-driven particles. *Physical Review Letters*, vol. 75, 1226-1229, 1995.
- [23] U. Krause, "Soziale dynamiken mit vielen interakteuren eine problemskizze," In *Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren*, pp. 37-51, 1997.
- [24] V. Blondel, J. M. Hendrickx and J. Tsitsiklis, "On Krause's consensus formation model with state-dependent connectivity," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2506-2517, 2009
- [25] V. Blondel, J. M. Hendrickx, A. Olshevsky and J. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," *IEEE Conf. Decision and Control*, pp. 2996-3000, 2005.
- [26] G. Tang and L. Guo, "Convergence of a class of multi-agent systems in probabilistic framework," *Journal of Systems Science and Complexity*, 20(2), pp. 173-197, 2007.

- [27] Z. Liu and L. Guo, “Synchronization of multi-agent systems without connectivity assumptions,” *Automatica*, vol. 45, no. 12, pp. 2744-2753, 2009.
- [28] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, “Continuous-time average-preserving opinion dynamics with opinion-dependent communications,” *SIAM Journal on Control and Optimization*, vol. 48, no. 8, pp. 5214-5240, 2010.
- [29] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale and V. Zdravkovic, “Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study,” *PNAS*, vol. 105, no. 4, pp. 1232-1237, 2008.
- [30] F. Xue and P. R. Kumar, “On the  $\theta$ -coverage and connectivity of large random networks,” *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2289-2299, 2006.
- [31] J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, “Distributed anonymous discrete function computation,” *IEEE Trans. Autom. Control*, vol. 56, no. 10, pp. 2276-2289, 2011.
- [32] B. M. Nejad, S. A. Attia and J. Raisch, “Max-consensus in a max-plus algebraic setting: The case of fixed communication topologies,” in *International Symposium on Information, Communication and Automation Technologies*, pp.1-7, 2009.
- [33] J. Cortés, “Finite-time convergent gradient flows with applications to network consensus,” *Automatica*, vol. 42, no.11, pp. 1993-2000, 2006.
- [34] J. Cortés, “Analysis and design of distributed algorithms for  $\chi$ -consensus,” *IEEE Conference on Decision and Control*, pp. 3363-3368, 2006.
- [35] F. Iutzeler, P. Ciblat, and J. Jakubowicz, “Analysis of max-consensus algorithms in wireless channels,” *IEEE Transactions on Signal Processing*, 2012 (submitted).
- [36] D. Varagnolo, G. Pilonetto, and L. Schenato, “Distributed statistical estimation of the number of nodes in sensor networks, in *IEEE Conference on Decision and Control*, 2011
- [37] R. Carli, F. Fagnani, A. Speranzon and S. Zampieri, “Communication constraints in the average consensus problem,” *Automatica*, vol.44, no. 3, pp. 671-684, 2008.
- [38] S. Boyd, A. Ghosh, B. Prabhakar and D. Shah, “Randomized gossip algorithms,” *IEEE Trans. Information Theory*, vol. 52, no. 6, pp. 2508-2530, 2006.

- [39] G. Shi, M. Johansson and K. H. Johansson, “When do gossip algorithms converge in finite time?” available at <http://arxiv.org/pdf/1206.0992v1.pdf>.